



Student Transcript and Academic Record Repository (STARR) Data Quality Technical Document

The purpose of this document is to provide the end user with specific business rules used in the creation of potential data quality reports that will be run by CEPI's data quality team during the current STARR collection. Please note that this document is not an exhaustive listing of all data quality checks. The listing below highlights the checks CEPI will be focusing on for the current STARR collection. Please refer to the CEPI [\(IHE\) Web page](#) for information and support documents regarding the STARR Collection.

NOTE: CEPI may periodically update this document.

Report Number	Report Title	Validation Rules	What to Look for?	Why?
ST1401*	Entry Dates greater than File Submission Date	{StudentDemographic}.[EntryDate] > {FileUpload}.[CreateDate]	Records where the {StudentDemographic}.[EntryDate] is after { FileUpload}.[CreateDate].	A File Upload Date cannot be a date greater than a student's Entry Date. Future data should not be reported.
ST1402	Entry Dates greater than Session Designator	{StudentDemographic}.[EntryDate] > {AcademicSession}.[SessionDesignator]	Records where the {StudentDemographic}.[EntryDate] is after {AcademicSession}.[SessionDesignator].	An Entry Date cannot be a date greater than a Session Designator at which a course is being reported. Future data should not be reported.
ST1403*	Student Level Code other than Non-degree Seeking Program submitted without a CIP Code.	{AcademicSession}.[StudentLevelCode] NE 20, 21, or 22 AND {AcademicProgram}.[ProgramCIPCode] IS MISSING	Records where a student was reported with {AcademicSession}.[StudentLevelCode] other than 20, 21, or 22 and {AcademicProgram}.[ProgramCIPCode] is missing.	Degree seeking students should be reported with Program CIP codes.
ST1405	Session Designator after July of Year Being Reported	{AcademicSession}.[SessionDesignator] > 201X-07	Records where the {AcademicSession}.[SessionDesignator] is after July of the current year being reported.	A Session Designator cannot be a date greater than July of the current year being reported. Future data should not be reported.
ST1420	Student's Entry Date	Prior{StudentDemographic}.[EntryDate]	Records where the	Once a student has officially

Report Number	Report Title	Validation Rules	What to Look for?	Why?
	changed from a previous STARR Collection	te] NOT = Current {StudentDemographic }.[EntryDate]	{StudentDemographic}. [EntryDate] differs from the previous year within your Institution.	entered the IHE, the first entry date should never change. The entry date should remain constant for the student at that IHE.
ST1422*	No Students Reported as Dual Enrolled, Other School Student, Concurrent Enrollment or Early Middle College	{AcademicSession}.[HighSchoolStudentId] is missing or not reported	Records where no students were reported as Dual Enrolled, Other Enrollment, or Early Middle College using the {AcademicSession}.[HighSchoolStudentId] field.	It is important to report this field for MI School Data reporting purposes as well as assisting colleges in submitting portions of the Section 209 and 245 Budget Transparency reporting when the IHE has students who participate in these programs.
ST1424	Academic Session Data not aligned with Academic Session Elements	(({AcademicSession}.[SessionName]= 'Fall' AND MONTH({AcademicSession}.[SessionDesignator]) NE 8,9,10,11, or 12) OR ({AcademicSession}.[SessionName]= 'Winter' AND MONTH({AcademicSession}.[SessionDesignator]) NE 1,2,3, or 4) OR ({AcademicSession}.[SessionName]= 'Spring' AND MONTH({AcademicSession}.[SessionDesignator]) NE 1,2,3, or 4) OR ({AcademicSession}.[SessionName]= 'Summer' AND	Records with {AcademicSession}.[SessionName] that do not fall within the range of the {AcademicSession}.[SessionDesignator] as recommended in the STARR manual	The month in the Session Designator does not fall within the range of the recommended guidelines in the STARR Manual.

Report Number	Report Title	Validation Rules	What to Look for?	Why?
		MONTH({AcademicSession}.[Session Designator]) NE 5,6,7 or 8)		
ST1425	High or low number of degree seeking students	(COUNT DISTINCT UICs ({AcademicSession}.[StudentLevelCode] = 20, 21, or 22))/ Total UICs LE 20% or GE 80%	Records where the ratio of UICs with {AcademicSession}.[StudentLevelCode] = 20,21,22 over total UICs is greater than 0.8 or less than 0.2	Having almost the entire student population with a student level code of degree seeking status may suggest an incomplete submission or an anomaly. On the other hand, a low ratio of students with a degree seeking status may also suggest an incomplete submission or an anomaly.
ST1501	Data Not Reported	COUNT(MONTH({AcademicSession}.[SessionDesignator]) IN 5,6 AND {AcademicSession}.[AcademicYearDesignator] = CURRENT) = 0	Count of {AcademicSession}.[SessionDesignator] during months 5 and 6 for CURRENT {AcademicSession}.[AcademicYearDesignator] = 0.	Summer session should have been reported in the current STARR Collection.
ST1503*	Same Academic Sessions Reported Multiple Times, No Primary	Count of {AcademicSession}.[SessionDesignator] NOT = Count distinct of {AcademicSession}.[SessionDesignator] AND duplicated {AcademicSession}.[SessionDesignator] does NOT contain {AcademicSession}.[PrimaryAcademicLevelId]	{AcademicSession}.[SessionDesignator] is reported multiple times for the same student and it is not indicated which {AcademicSession} is primary in {AcademicSession}.[PrimaryAcademicLevelId].	When the same Academic Session is reported multiple times for the same student, you must also indicate which Academic Session was primary for MI School Data reporting purposes.

Notes:

There are three windows in which STARR data quality will be run: pre-collection, mid-collection and post-collection.

- Pre-collection data quality analyzes prior year's data.

- For example: The pre-collection for STARR 2016 will analyze data from STARR 2015.
- If data are being compared to a previous year during the pre-collection window, the comparison will be the prior STARR collection to the STARR collection that occurred before it.
- Mid-collection data quality analyzes the current data that are being submitted up to the snapshot date.
 - The snapshot[^] date will be provided in the applicable data quality email(s).
 - If data are being compared to a previous year during the mid-collection window, the comparison will be the current STARR collection to the prior STARR collection.
- Post-collection data quality analyzes the current data after the STARR collection has been closed.
 - The comparisons utilize the current STARR Collection and/or prior STARR collection, similar to the mid-collection timeframe.

* These issues are NOT run during the pre-collection window.

** College Navigator's IPEDs code is located at: <https://nces.ed.gov/collegenavigator/>

[^] NCES CIP Codes are located at: <http://nces.ed.gov/ipeds/cipcode/browse.aspx?v=55>

[^] Snapshot is the capture of data at a point in time